Principal angles between subspaces and their tangents[☆]

Peizhen Zhu[a], Andrew V. Knyazev[a,b,*]

[a]Department of Mathematical and Statistical Sciences; University of Colorado Denver, P.O. Box 173364, Campus Box 170, Denver, CO 80217-3364, USA.
[b]Mitsubishi Electric Research Laboratories; 201 Broadway Cambridge, MA 02139

Abstract

Principal angles between subspaces (PABS) (also called canonical angles) serve as a classical tool in mathematics, statistics, and applications, e.g., data mining. Traditionally, PABS are introduced and used via their cosines. The tangents of PABS have attracted relatively less attention, but are important for analysis of convergence of subspace iterations for eigenvalue problems. We explicitly construct matrices, such that their singular values are equal to the tangents of PABS, using several approaches: orthonormal and non-orthonormal bases for subspaces, and orthogonal projectors.

Keywords: principal angles, canonical angles, singular value, projector
2000 MSC: 15A42, 15A60, 65F35

1. Introduction

The concept of principal angles between subspaces (PABS) is first introduced by Jordan [1] in 1875. Hotelling [2] defines PABS in the form of canonical correlations in statistics in 1936. Numerous researchers work on PABS; see, e.g., our pseudo-random choice of initial references [3, 4, 5, 6, 7, 8], out of tens of thousand of Internet links for the keywords “principal angles” and “canonical angles”.

☆September 5, 2012. The preliminary version is at `http://arxiv.org/abs/math/`. This material is based upon work partially supported by the National Science Foundation under Grant No. 1115734.

*Corresponding author.
Email addresses: `peizhen.zhu[at]ucdenver.edu` (Peizhen Zhu), `andrew.knyazev[at]ucdenver.edu` (Andrew V. Knyazev)
URL: `http://math.ucdenver.edu/~aknyazev/` (Andrew V. Knyazev)

In this paper, we first briefly review the concept and some important properties of PABS in Section 2. Traditionally, PABS are introduced and used via their sines and more commonly, because of the connection to canonical correlations, cosines. The tangents of PABS have attracted relatively less attention, despite of the celebrated work of Davis and Kahan [9], which includes several tangent-related theorems. Our interest to the tangents of PABS is motivated by their applications in theoretical analysis of convergence of subspace iterations for eigenvalue problems.

We review some previous work on the tangents of PABS in Section 3. The main goal of this work is explicitly constructing a family of matrices such that their singular values are equal to the tangents of PABS. We form these matrices using several different approaches: orthonormal bases for subspaces in Section 4, non-orthonormal bases in Section 5, and orthogonal projectors in Section 6.

2. Definition and basic properties of PABS

In this section, we remind the reader the concept of PABS and some fundamental properties of PABS. First, we recall that an acute angle between two unit vectors x and y, i.e., with $x^H x = y^H y = 1$, is defined as $\cos\theta(x, y) = |x^H y|$, where $0 \le \theta(x, y) \le \pi/2$.

The definition of an acute angle between two vectors can be recursively extended to PABS; see, e.g., [2, 10, 11].

Definition 2.1. *Let $\mathcal{X} \subset \mathbb{C}^n$ and $\mathcal{Y} \subset \mathbb{C}^n$ be subspaces with $\dim(\mathcal{X}) = p$ and $\dim(\mathcal{Y}) = q$. Let $m = \min(p, q)$. The principal angles*

$$\Theta(\mathcal{X}, \mathcal{Y}) = [\theta_1, \ldots, \theta_m]\,, \textit{where } \theta_k \in [0, \pi/2],\ k = 1, \ldots, m,$$

between $\mathcal{X}$ and $\mathcal{Y}$ are recursively defined by

$$s_k = \cos(\theta_k) = \max_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} |x^H y| = |x_k^H y_k|,$$

subject to

$$\|x\| = \|y\| = 1, \quad x^H x_i = 0, \quad y^H y_i = 0, \quad i = 1, \ldots, k-1.$$

The vectors $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_m\}$ are called the principal vectors.

An alternative definition of PABS, from [10, 11], is based on the singular value decomposition (SVD) and reproduced here as the following theorem.

Theorem 2.1. *Let the columns of matrices $X \in \mathbb{C}^{n\times p}$ and $Y \in \mathbb{C}^{n\times q}$ form orthonormal bases for the subspaces $\mathcal{X}$ and $\mathcal{Y}$, correspondingly. Let the SVD of X^HY be $U\Sigma V^H$, where U and V are unitary matrices and Σ is a $p \times q$ diagonal matrix with the real diagonal elements $s_1(X^HY), \ldots, s_m(X^HY)$ in nonincreasing order with $m = \min(p, q)$. Then*

$$\cos \Theta^{\uparrow}(\mathcal{X}, \mathcal{Y}) = S\left(X^HY\right) = \left[s_1\left(X^HY\right), \ldots, s_m\left(X^HY\right)\right],$$

where $\Theta^{\uparrow}(\mathcal{X}, \mathcal{Y})$ denotes the vector of principal angles between $\mathcal{X}$ and $\mathcal{Y}$ arranged in nondecreasing order and $S(A)$ denotes the vector of singular values of A. Moreover, the principal vectors associated with this pair of subspaces are given by the first m columns of XU and YV, correspondingly.

Theorem 2.1 implies that PABS are symmetric, i.e. $\Theta(\mathcal{X}, \mathcal{Y}) = \Theta(\mathcal{Y}, \mathcal{X})$, and unitarily invariant, i.e., $\Theta(U\mathcal{X}, U\mathcal{Y}) = \Theta(\mathcal{X}, \mathcal{Y})$ for any unitary matrix $U \in \mathbb{C}^{n\times n}$, since $(UX)^H(UY) = X^HY$. A number of other important properties of PABS have been established, for finite dimensional subspaces, e.g., in [4, 5, 6, 7, 8], and for infinite dimensional subspaces in [3, 12]. We list a few most useful properties of PABS below.

Property 2.1. [7] *In the notation of Theorem 2.1, let $p \geq q$ and let $[X, X_\perp]$ be a unitary matrix such that $S(X_\perp^HY) = [s_1(X_\perp^HY), \ldots, s_q(X_\perp^HY)]$ where $s_1(X_\perp^HY) \geq \cdots \geq s_q(X_\perp^HY)$. Then $s_k(X_\perp^HY) = \sin(\theta_{q+1-k})$, $k = 1, \ldots, q$.*

Relationships of principal angles between $\mathcal{X}$ and $\mathcal{Y}$, and between their orthogonal complements are thoroughly investigated in [4, 5, 12]. The nonzero principal angles between $\mathcal{X}$ and $\mathcal{Y}$ are the same as those between $\mathcal{X}^\perp$ and $\mathcal{Y}^\perp$. Similarly, the nonzero principal angles between $\mathcal{X}$ and $\mathcal{Y}^\perp$ are the same as those between $\mathcal{X}^\perp$ and $\mathcal{Y}$.

Property 2.2. [4, 5, 12] *Let the orthogonal complements of the subspaces $\mathcal{X}$ and $\mathcal{Y}$ be denoted by $\mathcal{X}^\perp$ and $\mathcal{Y}^\perp$, correspondingly. Then,*

1. $\left[\Theta^{\downarrow}(\mathcal{X}, \mathcal{Y}), 0, \ldots, 0\right] = \left[\Theta^{\downarrow}(\mathcal{X}^\perp, \mathcal{Y}^\perp), 0, \ldots, 0\right].$
2. $\left[\Theta^{\downarrow}(\mathcal{X}, \mathcal{Y}^\perp), 0, \ldots, 0\right] = \left[\Theta^{\downarrow}(\mathcal{X}^\perp, \mathcal{Y}), 0, \ldots, 0\right].$
3. $\left[\frac{\pi}{2}, \ldots, \frac{\pi}{2}, \Theta^{\downarrow}(\mathcal{X}, \mathcal{Y})\right] = \left[\frac{\pi}{2} - \Theta^{\uparrow}(\mathcal{X}, \mathcal{Y}^\perp), 0, \ldots, 0\right],$
   where there are $\max(\dim(\mathcal{X}) - \dim(\mathcal{Y}), 0)$ *additional $\pi/2$s on the left.*

Extra 0s at the end may need to be added on either side to match the sizes. Symbols $\downarrow$ and $\uparrow$ denote the components in vector arranged in nonincreasing and nondecreasing order, correspondingly.

These statements are widely used to discover new properties of PABS.

3. Tangents of PABS

The tangents of PABS serve as an important tool in numerical matrix analysis. For example, in [13, 14], the authors use the tangent of the largest principal angle derived from a norm of a specific matrix. In [6, p. 231-232] and [15, Theorem 2.4, p. 252] the tangents of PABS, related to singular values of a matrix—without an explicit matrix formulation—are used to analyze perturbations of invariant subspaces. In [6, 13, 14, 15], the two subspaces have the same dimensions.

Let the orthonormal columns of matrices X, $X_\perp$, and Y span the subspaces $\mathcal{X}$, the orthogonal complement $\mathcal{X}^\perp$ of $\mathcal{X}$, and $\mathcal{Y}$, correspondingly. According to Theorem 2.1 and Property 2.1, $\cos\Theta(\mathcal{X},\mathcal{Y}) = S(X^H Y)$ and $\sin\Theta(\mathcal{X},\mathcal{Y}) = S(X_\perp^H Y)$. The properties of sines and especially cosines of PABS are well studied; e.g., in [7, 10]. One could obtain the tangents of PABS by using the sines and the cosines of PABS directly, however, in some situations this does not reveal enough information about their properties.

In this work, we construct a family $\mathcal{F}$ of explicitly given matrices, such that the singular values of the matrix T are the tangents of PABS, where $T \in \mathcal{F}$. We form T in three different ways. First, we derive T as the product of matrices whose columns form the orthonormal bases of subspaces. Second, we present T as the product of matrices with non-orthonormal columns. Third, we form T as the product of orthogonal projections on subspaces. To better understand the matrix T, we provide a geometric interpretation of singular values of T and of the action of T as a linear operator.

Our constructions include the matrices used in [6, 13, 14, 15]. Furthermore, we consider the tangents of principal angles between two subspaces not only with the same dimensions, but also with different dimensions.

4. $\tan\Theta$ in terms of the orthonormal bases of subspaces

The sines and cosines of PABS are obtained from the singular values of explicitly given matrices, which motivates us to construct T as follows: if matrices X and Y have orthonormal columns and $X^H Y$ has full rank, let $T = X_\perp^H Y \left(X^H Y\right)^{-1}$, then $\tan\Theta(\mathcal{X},\mathcal{Y})$ could be equal to the positive singular values of T. We begin with an example in two dimensions in Figure 1.

$$\text{Let } X = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}, \quad X_\perp = \begin{pmatrix} -\sin\beta \\ \cos\beta \end{pmatrix}, \quad \text{and} \quad Y = \begin{pmatrix} \cos(\theta+\beta) \\ \sin(\theta+\beta) \end{pmatrix},$$

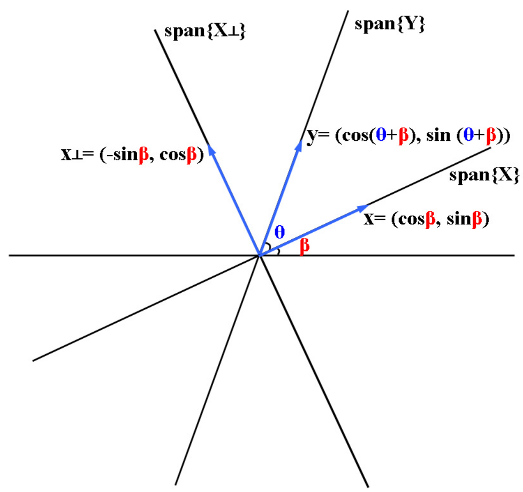

Figure 1: The PABS in 2D.

where $0 \leq \theta < \pi/2$ and $\beta \in [0, \pi/2]$. Then, $X^H Y = \cos\theta$ and $X_\perp^H Y = \sin\theta$. Obviously, $\tan\theta$ coincides with the singular value of $T = X_\perp^H Y \left(X^H Y\right)^{-1}$. If $\theta = \pi/2$, then the matrix $X^H Y$ is singular, which demonstrates that, in general, we need to use its Moore-Penrose Pseudoinverse to form our matrix $T = X_\perp^H Y \left(X^H Y\right)^\dagger$.

The Moore-Penrose Pseudoinverse is well known. For a matrix $A \in \mathbb{C}^{n\times m}$, its Moore-Penrose Pseudoinverse is the unique matrix $A^\dagger \in \mathbb{C}^{m\times n}$ satisfying

$$AA^\dagger A = A, \qquad A^\dagger A A^\dagger = A^\dagger, \qquad (AA^\dagger)^H = AA^\dagger, \qquad (A^\dagger A)^H = A^\dagger A.$$

Some properties of the Moore-Penrose Pseudoinverse are listed as follows.

- If $A = U\Sigma V^H$ is the SVD of A, then $A^\dagger = V\Sigma^\dagger U^H$.
- If A has full column rank and B has full row rank, then $(AB)^\dagger = B^\dagger A^\dagger$. However, this formula does not hold in general.
- $AA^\dagger$ is the orthogonal projector onto the range of A, and $A^\dagger A$ is the orthogonal projector onto the range of A^H.

For additional properties of the Moore-Penrose Pseudoinverse, we refer the reader to [6]. Now we are ready to prove that the intuition discussed above, suggesting to try $T = X_\perp^H Y \left(X^H Y\right)^\dagger$, is correct. Moreover, we cover the case of the subspaces $\mathcal{X}$ and $\mathcal{Y}$ having possibly different dimensions.

Theorem 4.1. *Let $X \in \mathbb{C}^{n\times p}$ have orthonormal columns and be arbitrarily completed to a unitary matrix $[X, X_\perp]$. Let the matrix $Y \in \mathbb{C}^{n\times q}$ be such that $Y^H Y = I$. Then the positive singular values, denoted by $S_+(T)$, of the matrix $T = X_\perp^H Y \left(X^H Y\right)^\dagger$ satisfy*

$$\tan \Theta(\mathcal{R}(X), \mathcal{R}(Y)) \;=\; [\infty, \ldots, \infty, S_+(T), 0, \ldots, 0],$$

where $\mathcal{R}(\cdot)$ denotes the matrix range.

Proof. Let $[\,Y \quad Y_\perp\,]$ be unitary, then $[\,X \quad X_\perp\,]^H[\,Y \quad Y_\perp\,]$ is unitary, such that

$$[\,X \quad X_\perp\,]^H[\,Y \quad Y_\perp\,] \;=\; \begin{array}{c c} & \begin{array}{cc} q & n-q \end{array} \\ \begin{array}{c} p \\ n-p \end{array} & \left[\begin{array}{c|c} X^H Y & X^H Y_\perp \\ \hline X_\perp^H Y & X_\perp^H Y_\perp \end{array}\right] \end{array}.$$

Applying the CS-decomposition (CSD), e.g., [16, 17, 18], we get

$$[\,X \quad X_\perp\,]^H[\,Y \quad Y_\perp\,] = \left[\begin{array}{c|c} X^H Y & X^H Y_\perp \\ \hline X_\perp^H Y & X_\perp^H Y_\perp \end{array}\right] = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} D \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^H,$$

where $U_1 \in \mathcal{U}(p), U_2 \in \mathcal{U}(n-p), V_1 \in \mathcal{U}(q)$, and $V_2 \in \mathcal{U}(n-q)$. The symbol $\mathcal{U}(k)$ denotes the family of unitary matrices of the size k. The matrix D has the following structure:

$$D = \begin{array}{c c} & \begin{array}{cccccc} r & s & q-r-s & n-p-q+r & s & p-r-s \end{array} \\ \begin{array}{c} r \\ s \\ p-r-s \\ n-p-q+r \\ s \\ q-r-s \end{array} & \left[\begin{array}{ccc|ccc} I & & & O_s^H & & \\ & C & & & S & \\ & & O_c & & & I \\ \hline O_s & & & -I & & \\ & S & & & -C & \\ & & I & & & O_c^H \end{array}\right] \end{array},$$

where $C = \operatorname{diag}(\cos(\theta_1), \ldots, \cos(\theta_s))$, and $S = \operatorname{diag}(\sin(\theta_1), \ldots, \sin(\theta_s))$ with $\theta_k \in (0, \pi/2)$ for $k = 1, \ldots, s$, which are the principal angles between the subspaces $\mathcal{R}(Y)$ and $\mathcal{R}(X)$. The matrices C and S could be not present. The matrices O_s and O_c are matrices of zeros and do not have to be square. I denotes the identity matrix. We may have different sizes of I in D.

Therefore,

$$\begin{aligned} T &= X_\perp^H Y \left(X^H Y\right)^\dagger = U_2 \begin{bmatrix} O_s & & \\ & S & \\ & & I \end{bmatrix} V_1^H V_1 \begin{bmatrix} I & & \\ & C & \\ & & O_c \end{bmatrix}^\dagger U_1^H \\ &= U_2 \begin{bmatrix} O_s & & \\ & SC^{-1} & \\ & & O_c^H \end{bmatrix} U_1^H. \end{aligned}$$

Hence, $S_+(T) = (\tan(\theta_1), \ldots, \tan(\theta_s))$, where $0 < \theta_1 \le \theta_2 \le \cdots \le \theta_s < \pi/2$.

From the matrix D, we obtain $S\left(X^H Y\right) = S\left(\operatorname{diag}(I, C, O_c)\right)$. According to Theorem 2.1, $\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = [0, \ldots, 0, \theta_1, \ldots, \theta_s, \pi/2, \ldots, \pi/2]$, where $\theta_k \in (0, \pi/2)$ for $k = 1, \ldots, s$, and there are $\min(q-r-s, p-r-s)$ additional $\pi/2$'s and r additional 0's on the right-hand side, which completes the proof. $\square$

Remark 4.2. *If more information is available about the subspaces $\mathcal{X}$, $\mathcal{Y}$ and their orthogonal complements, we can obtain the exact sizes of block matrices in the matrix D. Let us consider the decomposition of the space $\mathbb{C}^n$ into an orthogonal sum of five subspaces as in* [12, 19]*,*

$$\mathbb{C}^n = \mathfrak{M}_{00} \oplus \mathfrak{M}_{01} \oplus \mathfrak{M}_{10} \oplus \mathfrak{M}_{11} \oplus \mathfrak{M},$$

where $\mathfrak{M}_{00} = \mathcal{X} \cap \mathcal{Y}$, $\mathfrak{M}_{01} = \mathcal{X} \cap \mathcal{Y}^\perp$, $\mathfrak{M}_{10} = \mathcal{X}^\perp \cap \mathcal{Y}$, $\mathfrak{M}_{11} = \mathcal{X}^\perp \cap \mathcal{Y}^\perp$ ($\mathcal{X} = \mathcal{R}(X)$ and $\mathcal{Y} = \mathcal{R}(Y)$). Using Tables 1 and 2 in [12] *we get*

$$\begin{aligned} &\dim(\mathfrak{M}_{00}) = r, \\ &\dim(\mathfrak{M}_{10}) = q - r - s, \\ &\dim(\mathfrak{M}_{11}) = n - p - q + r, \\ &\dim(\mathfrak{M}_{01}) = p - r - s, \end{aligned}$$

where $\mathfrak{M} = \mathfrak{M}_\mathcal{X} \oplus \mathfrak{M}_{\mathcal{X}^\perp} = \mathfrak{M}_\mathcal{Y} \oplus \mathfrak{M}_{\mathcal{Y}^\perp}$ with

$$\begin{aligned} \mathfrak{M}_\mathcal{X} &= \mathcal{X} \cap \left(\mathfrak{M}_{00} \oplus \mathfrak{M}_{01}\right)^\perp, \\ \mathfrak{M}_{\mathcal{X}^\perp} &= \mathcal{X}^\perp \cap \left(\mathfrak{M}_{10} \oplus \mathfrak{M}_{11}\right)^\perp, \\ \mathfrak{M}_\mathcal{Y} &= \mathcal{Y} \cap \left(\mathfrak{M}_{00} \oplus \mathfrak{M}_{10}\right)^\perp, \\ \mathfrak{M}_{\mathcal{Y}^\perp} &= \mathcal{Y}^\perp \cap \left(\mathfrak{M}_{01} \oplus \mathfrak{M}_{11}\right)^\perp, \end{aligned}$$

and $s = \dim(\mathfrak{M}_{\mathcal{X}}) = \dim(\mathfrak{M}_{\mathcal{Y}}) = \dim(\mathfrak{M}_{\mathcal{X}^\perp}) = \dim(\mathfrak{M}_{\mathcal{Y}^\perp})$*. Thus, we can represent* D *in the following block form*

$$\begin{array}{c} \\ \dim(\mathfrak{M}_{00}) \\ \dim(\mathfrak{M})/2 \\ \dim(\mathfrak{M}_{01}) \\ \dim(\mathfrak{M}_{11}) \\ \dim(\mathfrak{M})/2 \\ \dim(\mathfrak{M}_{10}) \end{array}
\begin{array}{c} \begin{array}{cccccc} \dim(\mathfrak{M}_{00}) & \dim(\mathfrak{M})/2 & \dim(\mathfrak{M}_{10}) & \dim(\mathfrak{M}_{11}) & \dim(\mathfrak{M})/2 & \dim(\mathfrak{M}_{01}) \end{array} \\
\left[\begin{array}{ccc|ccc} I & & & O_s^H & & \\ & C & & & S & \\ & & O_c & & & I \\ \hline O_s & & & -I & & \\ & S & & & -C & \\ & & I & & & O_c^H \end{array}\right] \end{array}.$$

In addition, it is possible to permute the first q *columns or the last* $n-q$ *columns of* D*, or the first* p *rows or the last* $n-p$ *rows and to change the sign of any column or row to obtain the variants of the CSD.*

Remark 4.3. *From Remark 4.2, we see that some identity and zero matrices may be not present in the matrix* D*. To obtain more details for* D*, we have to deal with* D *in several cases based on either* $p \le q$ *or* $q \ge p$ *and either* $p+q \le n$ *or* $p+q \ge n$*. Since the angles are symmetric such that* $\Theta(\mathcal{X},\mathcal{Y}) = \Theta(\mathcal{Y},\mathcal{X})$*, it is enough to discuss one case of either* $p \le q$ *or* $q \le p$*. Let us assume* $q \le p$ *which implies* $\dim(\mathfrak{M}_{01}) - \dim(\mathfrak{M}_{10}) = p - q$*.*

Case 1*: assuming* $p+q \le n$*, we have (also see* [16, 20]*),*

$$D = \left[\begin{array}{c|ccc} C_0 & O & S_0 & O \\ O_{(p-q)\times q} & O & O & I_{p-q} \\ \hline -S_0 & O & C_0 & O \\ O_{(n-p-q)\times q} & I_{n-p-q} & O & O \end{array}\right],$$

where $C_0 = \operatorname{diag}(\cos(\theta_1), \ldots, \cos(\theta_q))$ *and* $S_0 = \operatorname{diag}(\sin(\theta_1), \ldots, \sin(\theta_q))$ *with* $\theta_k \in [0, \pi/2]$ *for* $k = 1, \ldots, q$*.*

Case 2 *: assuming* $p+q \ge n$*, we have*

$$D = \left[\begin{array}{cc|cc} C_0 & O_{(n-p)\times(p+q-n)} & S_0 & O \\ O_{(p+q-n)\times(n-p)} & I_{p+q-n} & O & O \\ O_{(p-q)\times(n-p)} & O_{(p-q)\times(p+q-n)} & O & I_{p-q} \\ \hline -S_0 & O_{(n-p)\times(p+q-n)} & C_0 & O \end{array}\right],$$

where $C_0 = \operatorname{diag}(\cos(\theta_1), \ldots, \cos(\theta_{n-p}))$ *and* $S_0 = \operatorname{diag}(\sin(\theta_1), \ldots, \sin(\theta_{n-p}))$ *with* $\theta_k \in [0, \pi/2]$ *for* $k = 1, \ldots, n-p$*.*

Due to the fact that the angles are symmetric, i.e., $\Theta(\mathcal{X},\mathcal{Y}) = \Theta(\mathcal{Y},\mathcal{X})$, the matrix T in Theorem 4.1 could be substituted with $Y_\perp^H X(Y^H X)^\dagger$. Moreover, the nonzero angles between the subspaces $\mathcal{X}$ and $\mathcal{Y}$ are the same as those between the subspaces $\mathcal{X}^\perp$ and $\mathcal{Y}^\perp$. Hence, T can be presented as $X^H Y_\perp (X_\perp^H Y_\perp)^\dagger$ and $Y^H X_\perp (Y_\perp^H X_\perp)^\dagger$. Furthermore, for any matrix T, we have $S(T) = S(T^H)$, which implies that all conjugate transposes of T are valid. Let $\mathcal{F}$ denote a family of matrices, such that the singular values of the matrix in $\mathcal{F}$ are the tangents of PABS. To sum up, any of the formulas for T in $\mathcal{F}$ in the first column of Table 1 can be used in Theorem 4.1.

Table 1: Different matrices in $\mathcal{F}$: matrix $T \in \mathcal{F}$ using orthonormal bases.

$X_\perp^H Y \left(X^H Y\right)^\dagger$	$P_{\mathcal{X}^\perp} Y \left(X^H Y\right)^\dagger$
$Y_\perp^H X (Y^H X)^\dagger$	$P_{\mathcal{Y}^\perp} X (Y^H X)^\dagger$
$X^H Y_\perp (X_\perp^H Y_\perp)^\dagger$	$P_{\mathcal{X}} Y_\perp (X_\perp^H Y_\perp)^\dagger$
$Y^H X_\perp (Y_\perp^H X_\perp)^\dagger$	$P_{\mathcal{Y}} X_\perp (Y_\perp^H X_\perp)^\dagger$
$(Y^H X)^\dagger Y^H X_\perp$	$(Y^H X)^\dagger Y^H P_{\mathcal{X}^\perp}$
$(X^H Y)^\dagger X^H Y_\perp$	$(X^H Y)^\dagger X^H P_{\mathcal{Y}^\perp}$
$(Y_\perp^H X_\perp)^\dagger Y_\perp^H X$	$(Y_\perp^H X_\perp)^\dagger Y_\perp^H P_{\mathcal{X}}$
$(X_\perp^H Y_\perp)^\dagger X_\perp^H Y$	$(X_\perp^H Y_\perp)^\dagger X_\perp^H P_{\mathcal{Y}}$

Using the fact that the singular values are invariant under unitary multiplications, we can also use $P_{\mathcal{X}^\perp} Y (X^H Y)^\dagger$ for T in Theorem 4.1, where $P_{\mathcal{X}^\perp}$ is an orthogonal projector onto the subspace $\mathcal{X}^\perp$. Thus, we can use T as in the second column in Table 1, where $P_{\mathcal{X}}$, $P_{\mathcal{Y}}$, and $P_{\mathcal{Y}^\perp}$ denote the orthogonal projectors onto the subspace $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Y}^\perp$, correspondingly.

Remark 4.4. *From the proof of Theorem 4.1, we immediately derive that* $X_\perp^H Y \left(X^H Y\right)^\dagger = -(Y_\perp^H X_\perp)^\dagger Y_\perp^H X$, *and* $Y_\perp^H X (Y^H X)^\dagger = -(X_\perp^H Y_\perp)^\dagger X_\perp^H Y$.

Remark 4.5. *Let* $X_\perp^H Y \left(X^H Y\right)^\dagger = U_2 \Sigma_1 U_1$ *and* $Y_\perp^H X (Y^H X)^\dagger = V_2 \Sigma_2 V_1$, *be the SVDs, where* U_i *and* V_i $(i = 1, 2)$ *are unitary matrices and* Σ_i $(i = 1, 2)$ *are diagonal matrices. Let the diagonal elements of* Σ_1 *and* Σ_2 *be in the same order. Then the principal vectors associated with the pair of subspaces* $\mathcal{X}$ *and* $\mathcal{Y}$ *are given by the columns of* XU_1 *and* YV_1.

5. $\tan\Theta$ in terms of the bases of subspaces

In the previous section, we use matrices X and Y with orthonormal columns to formulate T. It turns out that we can choose a non-orthonormal

basis for one of the subspaces $\mathcal{X}$ or $\mathcal{Y}$.

Theorem 5.1. *Let* $[X, X_\perp]$ *be a unitary matrix with* $X \in \mathbb{C}^{n\times p}$. *Let* $Y \in \mathbb{C}^{n\times q}$ *and* $\operatorname{rank}(Y) = \operatorname{rank}\left(X^H Y\right) \le p$. *Then the positive singular values of the matrix* $T = X_\perp^H Y \left(X^H Y\right)^\dagger$ *satisfy* $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = [S_+(T), 0, \ldots, 0]$.

Proof. Let the rank of Y be r, thus $r \le p$. Let the SVD of Y be $U\Sigma V^H$, where U is an $n \times n$ unitary matrix; Σ is an $n \times q$ real diagonal matrix with diagonal entries $s_1, \ldots, s_r, 0, \ldots, 0$ ordered by decreasing magnitude such that $s_1 \ge s_2 \ge \cdots \ge s_r > 0$, and V is a $q\times q$ unitary matrix. Since rank(Y)=$r \le q$, we can compute a reduced SVD such that $Y = U_r \Sigma_r V_r^H$. Only the r column vectors of U and the r row vectors of V^H, corresponding to nonzero singular values are used, which means that Σ_r is an $r \times r$ invertible diagonal matrix. Based on the fact that the left singular vectors corresponding to the non-zero singular values of Y span the range of Y, we have

$$\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = \tan\Theta(\mathcal{R}(X), \mathcal{R}(U_r)).$$

Since $\operatorname{rank}(X^H Y) = r$, we have $\operatorname{rank}(X^H U_r \Sigma_r) = \operatorname{rank}(X^H U_r) = r$. Let $T_1 = X_\perp^H U_r \left(X^H U_r\right)^\dagger$. According to Theorem 4.1, it follows that

$$\tan\Theta(\mathcal{R}(X), \mathcal{R}(U_r)) = [S_+(T_1), 0, \ldots, 0].$$

It is worth noting that the angles between $\mathcal{R}(X)$ and $\mathcal{R}(U_r)$ are in $[0, \pi/2)$, since $X^H U_r$ is full rank.

Our task is now to show that $T_1 = T$. By direct computation, we have

$$\begin{aligned} T &= X_\perp^H Y \left(X^H Y\right)^\dagger \\ &= X_\perp^H U_r \Sigma_r V_r^H \left(X^H U_r \Sigma_r V_r^H\right)^\dagger \\ &= X_\perp^H U_r \Sigma_r V_r^H \left(V_r^H\right)^\dagger \left(X^H U_r \Sigma_r\right)^\dagger \\ &= X_\perp^H U_r \Sigma_r \left(X^H U_r \Sigma_r\right)^\dagger \\ &= X_\perp^H U_r \Sigma_r \Sigma_r^\dagger \left(X^H U_r\right)^\dagger \\ &= X_\perp^H U_r \left(X^H U_r\right)^\dagger. \end{aligned}$$

In two identities above we use the fact that if a matrix A is of full column rank, and a matrix B is of full row rank, then $(AB)^\dagger = B^\dagger A^\dagger$. Hence, $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = [S_+(T), 0, \ldots, 0]$ which completes the proof. $\square$

Remark 5.2. *Let X^HY have full rank. Let us define $Z = X + X_\perp T$ following* [6, p.231-232] *and* [15, Theorem 2.4, p.252]. *Since $\left(X^HY\right)^\dagger X^HY = I$, the following identities $Y = P_{\mathcal{X}}Y + P_{\mathcal{X}^\perp}Y = XX^HY + X_\perp X_\perp^H Y = ZX^HY$ imply that $\mathcal{R}(Y) \subseteq \mathcal{R}(Z)$. By direct calculation, we obtain that*

$$X^HZ = X^H(X + X_\perp T) = I \text{ and } Z^HZ = (X + X_\perp T)^H(X + X_\perp T) = I + T^HT.$$

Thus, $X^HZ\left(Z^HZ\right)^{-1/2} = (I + T^HT)^{-1/2}$ is Hermitian positive definite. The matrix $Z(Z^HZ)^{-1/2}$ by construction has orthonormal columns which span the space $\mathcal{Z}$. Moreover, we observe that

$$S^\uparrow\left(X^HZ(Z^HZ)^{-1/2}\right) = \Lambda^\uparrow\left((I + T^HT)^{-1/2}\right) = \left(1 + S^2\left(T\right)\right)^{-1/2}.$$

Therefore, $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Z)) = [S_+(T), 0, \ldots, 0]$ and $\dim(\mathcal{X}) = \dim(\mathcal{Z})$. From Theorem 5.1, we have $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) \subset \tan\Theta(\mathcal{R}(X), \mathcal{R}(Z))$. It is worth noting that $\Theta(\mathcal{R}(X), \mathcal{R}(Y))$ and $\Theta(\mathcal{R}(X), \mathcal{R}(Z))$ in $(0, \pi/2)$ are the same. For the case $p = q$, we have that $\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = \Theta(\mathcal{R}(X), \mathcal{R}(Z))$. This gives us an explicit matrix expression $X_\perp^HY\left(X^HY\right)^{-1}$ for the matrix P in [6, p.231-232] *and* [15, Theorem 2.4, p.252].

Remark 5.3. *The condition $\operatorname{rank}(Y) = \operatorname{rank}\left(X^HY\right) \leq p$ is necessary in Theorem 5.1. For example, let*

$$X = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad X_\perp = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad Y = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then, we have $X^HY = \begin{pmatrix} 1 & 1 \end{pmatrix}$ and $(X^HY)^\dagger = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}^H$. Thus, $T = X_\perp^HY\left(X^HY\right)^\dagger = \begin{pmatrix} 1/2 & 0 \end{pmatrix}^H$, and $s(T) = 1/2$. On the other hand, we obtain that $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = 0$. From this example, we see that the result in Theorem 5.1 does not hold for the case $\operatorname{rank}(Y) = \operatorname{rank}\left(X^HY\right) > p$.

Corollary 5.1. *Using the notation of Theorem 5.1, let Y be full rank and $p = q$. Let $P_{\mathcal{X}^\perp}$ be an orthogonal projection onto the subspace $\mathcal{R}(X_\perp)$. Then $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = S\left(P_{\mathcal{X}^\perp}Y\left(X^HY\right)^{-1}\right)$.*

Proof. Since the singular values are invariant under unitary transform, it is easy to obtain $\left[S\left(X_\perp^HY\left(X^HY\right)^{-1}\right), 0, \ldots, 0\right] = S\left(P_{\mathcal{X}^\perp}Y\left(X^HY\right)^{-1}\right)$. Moreover, the number of the singular values of $P_{\mathcal{X}^\perp}Y\left(X^HY\right)^{-1}$ is p, hence we obtain the result as stated. $\square$

Above we construct T explicitly in terms of bases of subspaces $\mathcal{X}$ and $\mathcal{Y}$. Next, we provide an alternative construction by adopting only the triangular matrix from the QR factorization of a basis of the subspace $\mathcal{X}+\mathcal{Y}$.

Corollary 5.2. *Let $X \in \mathbb{C}^{n\times p}$ and $Y \in \mathbb{C}^{n\times q}$ be matrices of full rank where $q \le p$. Let the triangular part R of the reduced QR factorization of the matrix $L = [X, Y]$ be*

$$R = \begin{bmatrix} R_{11} & R_{12} \\ O & R_{22} \end{bmatrix},$$

where R_{12} is full rank. Then the positive singular values of the matrix $T = R_{22}(R_{12})^{\dagger}$satisfy $\tan\Theta(\mathcal{R}(X), \mathcal{R}(Y)) = [S_{+}(T), 0, \ldots, 0]$.

Proof. Let the reduced QR factorization of L be

$$\begin{bmatrix} X & Y \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ O & R_{22} \end{bmatrix},$$

where the columns of $[\ Q_1 \quad Q_2\ ]$ are orthonormal, and $\mathcal{R}(Q_1) = \mathcal{X}$. We directly obtain $Y = Q_1R_{12} + Q_2R_{22}$. Multiplying by Q_1^H on both sides of the above equality for Y, we get $R_{12} = Q_1^HY$. Similarly, multiplying by Q_2^H we have $R_{22} = Q_2^HY$. Combining these equalities with Theorem 5.1 completes the proof. □

6. tan Θ in terms of the orthogonal projections onto subspaces

In this section, we present T using orthogonal projectors only. Before we proceed, we need a lemma and a corollary.

Lemma 6.1. [6] *If U and V are unitary matrices, then for any matrix A we have $(UAV)^{\dagger} = V^HA^{\dagger}U^H$.*

Corollary 6.1. *Let $A \in \mathbb{C}^{n\times n}$ be a block matrix, such that*

$$A = \begin{bmatrix} D & O_{12} \\ O_{21} & O_{22} \end{bmatrix},$$

where D is a $p \times q$ matrix and O_{12}, O_{21}, and O_{22} are zero matrices. Then

$$A^{\dagger} = \begin{bmatrix} D^{\dagger} & O_{21}^H \\ O_{12}^H & O_{22}^H \end{bmatrix}.$$

Proof. Let $\mathcal{U}(n)$ denote a family of unitary matrices of size n by n. Let the SVD of D be $U_1\Sigma V_1^H$, where Σ is a $p\times q$ diagonal matrix and $U_1\in\mathcal{U}(p)$, $V_1\in\mathcal{U}(q)$. Then there exist unitary matrices $U_2\in\mathcal{U}(n-p)$ and $V_2\in\mathcal{U}(n-q)$, such that

$$A=\begin{bmatrix} U_1 & \\ & U_2\end{bmatrix}\begin{bmatrix}\Sigma & O_{12}\\ O_{21} & O_{22}\end{bmatrix}\begin{bmatrix}V_1^H & \\ & V_2^H\end{bmatrix}.$$

Thus,

$$\begin{aligned} A^\dagger &= \begin{bmatrix} V_1 & \\ & V_2\end{bmatrix}\begin{bmatrix}\Sigma & O_{12}\\ O_{21} & O_{22}\end{bmatrix}^\dagger\begin{bmatrix}U_1^H & \\ & U_2^H\end{bmatrix}\\ &= \begin{bmatrix} V_1 & \\ & V_2\end{bmatrix}\begin{bmatrix}\Sigma^\dagger & O_{21}^H\\ O_{12}^H & O_{22}^H\end{bmatrix}\begin{bmatrix}U_1^H & \\ & U_2^H\end{bmatrix}.\end{aligned}$$

Since $D^\dagger=V_1\Sigma^\dagger U_1^H$, we have $A^\dagger=\begin{bmatrix}D^\dagger & O_{21}^H\\ O_{12}^H & O_{22}^H\end{bmatrix}$ as desired. □

Theorem 6.2. *Let $P_{\mathcal{X}}$, $P_{\mathcal{X}^\perp}$ and $P_{\mathcal{Y}}$ be orthogonal projectors onto the subspaces $\mathcal{X}$, $\mathcal{X}^\perp$ and $\mathcal{Y}$, correspondingly. Then the positive singular values $S_+(T)$ of the matrix $T=P_{\mathcal{X}^\perp}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger$ satisfy*

$$\tan\Theta(\mathcal{X},\mathcal{Y})=[\infty,\ldots,\infty,S_+(T),0,\ldots,0].$$

Proof. Let the matrices $[\,X\quad X_\perp\,]$ and $[\,Y\quad Y_\perp\,]$ be unitary, where $\mathcal{R}(X)=\mathcal{X}$ and $\mathcal{R}(Y)=\mathcal{Y}$. By direct calculation, we obtain

$$\begin{bmatrix}X^H\\ X_\perp^H\end{bmatrix}P_{\mathcal{X}}P_{\mathcal{Y}}[\,Y\quad Y_\perp\,]=\begin{bmatrix}X^HY & O\\ O & O\end{bmatrix}.$$

The matrix $P_{\mathcal{X}}P_{\mathcal{Y}}$ can be written as

$$P_{\mathcal{X}}P_{\mathcal{Y}} = [\,X\quad X_\perp\,]\begin{bmatrix}X^HY & O\\ O & O\end{bmatrix}\begin{bmatrix}Y^H\\ Y_\perp^H\end{bmatrix}. \tag{1}$$

From Lemma 6.1 and Corollary 6.1, we have

$$(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = [\,Y\quad Y_\perp\,]\begin{bmatrix}(X^HY)^\dagger & O\\ O & O\end{bmatrix}\begin{bmatrix}X^H\\ X_\perp^H\end{bmatrix}. \tag{2}$$

In a similar way,

$$P_{\mathcal{X}^\perp}P_{\mathcal{Y}} = [\,X_\perp\quad X\,]\begin{bmatrix}X_\perp^HY & O\\ O & O\end{bmatrix}\begin{bmatrix}Y^H\\ Y_\perp^H\end{bmatrix}. \tag{3}$$

Let $T = P_{\mathcal{X}^\perp} P_{\mathcal{Y}} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$ and $B = X_\perp^H Y \left(X^H Y\right)^\dagger$, so

$$\begin{aligned} T &= \left[\begin{array}{cc} X_\perp & X \end{array}\right] \left[\begin{array}{cc} X_\perp^H Y & O \\ O & O \end{array}\right] \left[\begin{array}{cc} (X^H Y)^\dagger & O \\ O & O \end{array}\right] \left[\begin{array}{c} X^H \\ X_\perp^H \end{array}\right] \\ &= \left[\begin{array}{cc} X_\perp & X \end{array}\right] \left[\begin{array}{cc} B & O \\ O & O \end{array}\right] \left[\begin{array}{c} X^H \\ X_\perp^H \end{array}\right]. \end{aligned}$$

The singular values are invariant under unitary multiplications, therefore $S_+(T) = S_+(B)$.

Finally, we show that $P_{\mathcal{X}^\perp} P_{\mathcal{Y}} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger = P_{\mathcal{X}^\perp} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$. The null space of $P_{\mathcal{X}} P_{\mathcal{Y}}$ is the orthogonal sum $\mathcal{Y}^\perp \oplus \left(\mathcal{Y} \cap \mathcal{X}^\perp\right)$. The range of $(P_{\mathcal{X}} P_{\mathcal{Y}})^\dagger$ is thus $\mathcal{Y} \cap \left(\mathcal{Y} \cap \mathcal{X}^\perp\right)^\perp$, which is the orthogonal complement of the null space of $P_{\mathcal{X}} P_{\mathcal{Y}}$. Moreover, $(P_{\mathcal{X}} P_{\mathcal{Y}})^\dagger$ is an oblique projector, because $(P_{\mathcal{X}} P_{\mathcal{Y}})^\dagger = \left((P_{\mathcal{X}} P_{\mathcal{Y}})^\dagger\right)^2$ which can be obtained by direct calculation using equality (2). The oblique projector $(P_{\mathcal{X}} P_{\mathcal{Y}})^\dagger$ projects onto $\mathcal{Y}$ along $\mathcal{X}^\perp$. Thus, we have $P_{\mathcal{Y}} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger = \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$. Consequently, we can simplify $T = P_{\mathcal{X}^\perp} P_{\mathcal{Y}} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$ as $T = P_{\mathcal{X}^\perp} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$. $\square$

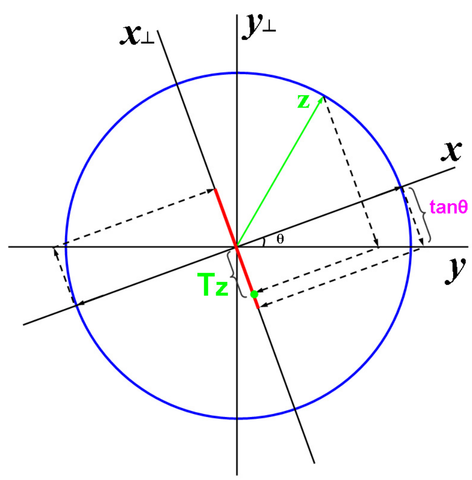

Figure 2: Geometrical meaning of $T = P_{\mathcal{X}^\perp} \left(P_{\mathcal{X}} P_{\mathcal{Y}}\right)^\dagger$.

To gain geometrical insight of T, in Figure 2 we choose a generic unit vector z. We project z onto $\mathcal{Y}$ along $\mathcal{X}^\perp$, then project onto the subspace $\mathcal{X}^\perp$ which is interpreted as Tz. The red segment in the graph is the image of T under all unit vectors. It is straightforward to see that $s(T) = \|T\| = \tan(\theta)$.

Using Property 2.2 and the fact that the principal angles are symmetric with respect to the subspaces $\mathcal{X}$ and $\mathcal{Y}$, the expressions $P_{\mathcal{Y}^\perp}(P_{\mathcal{Y}}P_{\mathcal{X}})^\dagger$ and $P_{\mathcal{X}}(P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger$ can also be used in Theorem 6.2.

Remark 6.3. *According to Remark 4.4 and Theorem 6.2, we have*

$$P_{\mathcal{X}}(P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger = P_{\mathcal{X}}P_{\mathcal{Y}^\perp}(P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger = -\left(P_{\mathcal{X}^\perp}\left(P_{\mathcal{X}}P_{\mathcal{Y}}\right)^\dagger\right)^H.$$

Furthermore, by the geometrical properties of the oblique projector $(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger$*, it follows that* $(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = P_{\mathcal{Y}}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger P_{\mathcal{X}}$*. Therefore, we obtain*

$$P_{\mathcal{X}^\perp}\left(P_{\mathcal{X}}P_{\mathcal{Y}}\right)^\dagger = P_{\mathcal{X}^\perp}P_{\mathcal{Y}}\left(P_{\mathcal{X}}P_{\mathcal{Y}}\right)^\dagger = (P_{\mathcal{Y}} - P_{\mathcal{X}})(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger.$$

Similarly, we have

$$P_{\mathcal{Y}^\perp}(P_{\mathcal{Y}}P_{\mathcal{X}})^\dagger = P_{\mathcal{Y}^\perp}P_{\mathcal{X}}(P_{\mathcal{Y}}P_{\mathcal{X}})^\dagger = (P_{\mathcal{X}} - P_{\mathcal{Y}})(P_{\mathcal{Y}}P_{\mathcal{X}})^\dagger = -\left(P_{\mathcal{Y}}(P_{\mathcal{Y}^\perp}P_{\mathcal{X}^\perp})^\dagger\right)^H.$$

From equalities above, we see that there are several different expressions for T *in Theorem 6.2.*

Theorem 6.4. *For subspaces* $\mathcal{X}$ *and* $\mathcal{Y}$*,*

1. *If* $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$*, we have* $P_{\mathcal{X}}P_{\mathcal{Y}}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = P_{\mathcal{X}}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = P_{\mathcal{X}}$*.*
2. *If* $\dim(\mathcal{X}) \geq \dim(\mathcal{Y})$*, we have* $(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger P_{\mathcal{X}}P_{\mathcal{Y}} = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger P_{\mathcal{Y}} = P_{\mathcal{Y}}$*.*

Proof. Using the fact that $X^HY(X^HY)^\dagger = I$ for $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$ and combining identities (1) and (2), we obtain the first statement. Identities (1) and (2), together with the fact that $(X^HY)^\dagger X^HY = I$ for $\dim(\mathcal{X}) \geq \dim(\mathcal{Y})$, imply the second statement. □

Remark 6.5. *From Theorem 6.4, for the case* $\dim(\mathcal{X}) \leq \dim(\mathcal{Y})$ *we have*

$$P_{\mathcal{X}^\perp}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger - P_{\mathcal{X}}(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger - P_{\mathcal{X}}.$$

Since the angles are symmetric, using the second statement in Theorem 6.4 we have $(P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger P_{\mathcal{Y}} = (P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger - P_{\mathcal{Y}^\perp}$*.*

On the other hand, for the case $\dim(\mathcal{X}) \geq \dim(\mathcal{Y})$ *we obtain*

$$(P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger P_{\mathcal{Y}^\perp} = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger - P_{\mathcal{Y}} \text{ and } (P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger P_{\mathcal{X}} = (P_{\mathcal{X}^\perp}P_{\mathcal{Y}^\perp})^\dagger - P_{\mathcal{X}^\perp}.$$

To sum up, the following formulas for T *in Table 2 can also be used in Theorem 6.2. An alternative proof for* $T = (P_{\mathcal{X}}P_{\mathcal{Y}})^\dagger - P_{\mathcal{Y}}$ *is provided by Drmač in* [14] *for the particular case* $\dim(\mathcal{X}) = \dim(\mathcal{Y})$*.*

Table 2: Different formulas for T: left for $\dim(\mathcal{X}) \le \dim(\mathcal{Y})$; right for $\dim(\mathcal{X}) \ge \dim(\mathcal{Y})$.

$(P_{\mathcal{X}}P_{\mathcal{Y}})^{\dagger} - P_{\mathcal{X}}$	$(P_{\mathcal{X}}P_{\mathcal{Y}})^{\dagger} - P_{\mathcal{Y}}$
$(P_{\mathcal{X}^{\perp}}P_{\mathcal{Y}^{\perp}})^{\dagger} - P_{\mathcal{Y}^{\perp}}$	$(P_{\mathcal{X}^{\perp}}P_{\mathcal{Y}^{\perp}})^{\dagger} - P_{\mathcal{X}^{\perp}}$

Remark 6.6. *Finally, we note that our choice of the space $\mathcal{H} = \mathbb{C}^n$ may appear natural to the reader familiar with the matrix theory, but in fact is somewhat misleading. The principal angles (and the corresponding principal vectors) between the subspaces $\mathcal{X} \subset \mathcal{H}$ and $\mathcal{Y} \subset \mathcal{H}$ are exactly the same as those between the subspaces $\mathcal{X} \subset \mathcal{X} + \mathcal{Y}$ and $\mathcal{Y} \subset \mathcal{X} + \mathcal{Y}$, i.e., we can reduce the space $\mathcal{H}$ to the space $\mathcal{X} + \mathcal{Y} \subset \mathcal{H}$ without changing PABS.*

This reduction changes the definition of the subspaces $\mathcal{X}^{\perp}$ and $\mathcal{Y}^{\perp}$ and, thus, of the matrices $X_{\perp}$ and $Y_{\perp}$ that column-span the subspaces $\mathcal{X}^{\perp}$ and $\mathcal{Y}^{\perp}$. All our statements that use the subspaces $\mathcal{X}^{\perp}$ and $\mathcal{Y}^{\perp}$ or the matrices $X_{\perp}$ and $Y_{\perp}$ therefore have their new analogs, if the space $\mathcal{X} + \mathcal{Y}$ substitutes for $\mathcal{H}$. The formulas $P_{\mathcal{X}^{\perp}} = I - P_{\mathcal{X}^{\perp}}$ and $P_{\mathcal{Y}^{\perp}} = I - P_{\mathcal{Y}^{\perp}}$, look the same, but the identity operators are different in the spaces $\mathcal{X} + \mathcal{Y}$ and $\mathcal{H}$.

References

[1] C. Jordan, Essai sur la géométrie á n dimensions, Amer. Math. Monthly 3 (1875) 103–174.

[2] H. Hotelling, Relations between two sets of variates, Biometrika 28 (1936) 321–377.

[3] F. Deutsch, The angle between subspaces of a Hilbert space, in: Approximation theory, wavelets and applications (Maratea, 1994), volume 454 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, Kluwer Acad. Publ., Dordrecht, 1995, pp. 107–130.

[4] I. C. F. Ipsen, C. D. Meyer, The angle between complementary subspaces, Amer. Math. Monthly 102 (1995) 904–911.

[5] A. V. Knyazev, M. E. Argentati, Majorization for changes in angles between subspaces, Ritz values, and graph Laplacian spectra, SIAM J. Matrix Anal. Appl. 29 (2006/07) 15–32 (electronic).

[6] G. W. Stewart, J. Sun, Matrix perturbation theory, Computer Science and Scientific Computing, Academic Press Inc., Boston, MA, 1990.

[7] J. Sun, Perturbation of angles between linear subspaces, J. Comput. Math. 5 (1987) 58–61.

[8] P. Wedin, On angles between subspaces of a finite dimensional inner product space, in: B. Kgstrm, A. Ruhe (Eds.), Matrix Pencils, volume 973 of *Lecture Notes in Mathematics*, Springer Berlin Heidelberg, 1983, pp. 263–285.

[9] C. Davis, W. M. Kahan, The rotation of eigenvectors by a perturbation. III, SIAM J. Numer. Anal. 7 (1970) 1–46.

[10] Ȧ. Björck, G. Golub, Numerical methods for computing angles between linear subspaces, Math. Comp. 27 (1973) 579–594.

[11] G. H. Golub, H. Zha, The canonical correlations of matrix pairs and their numerical computation, in: Linear algebra for signal processing (Minneapolis, MN, 1992), volume 69 of *IMA Vol. Math. Appl.*, Springer, New York, 1995, pp. 27–49.

[12] A. V. Knyazev, A. Jujunashvili, M. Argentati, Angles between infinite dimensional subspaces with applications to the Rayleigh-Ritz and alternating projectors methods, J. Funct. Anal. 259 (2010) 1323–1345.

[13] N. Bosner, Z. Drmač, Subspace gap residuals for Rayleigh-Ritz approximations, SIAM J. Matrix Anal. Appl. 31 (2009) 54–67.

[14] Z. Drmač, On relative residual bounds for the eigenvalues of a Hermitian matrix, Linear Algebra Appl. 244 (1996) 155–163.

[15] G. W. Stewart, Matrix algorithms. Vol. II, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Eigensystems.

[16] A. Edelman, B. D. Sutton, The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems, Found. Comput. Math. 8 (2008) 259–285.

[17] C. C. Paige, M. A. Saunders, Towards a generalized singular value decomposition, SIAM J. Numer. Anal. 18 (1981) 398–405.

[18] C. C. Paige, M. Wei, History and generality of the CS decomposition, Linear Algebra Appl. 208/209 (1994) 303–326.

[19] P. R. Halmos, Two subspaces, Trans. Amer. Math. Soc. 144 (1969) 381–389.

[20] B. D. Sutton, Computing the complete CS decomposition, Numer. Algorithms 50 (2009) 33–65.